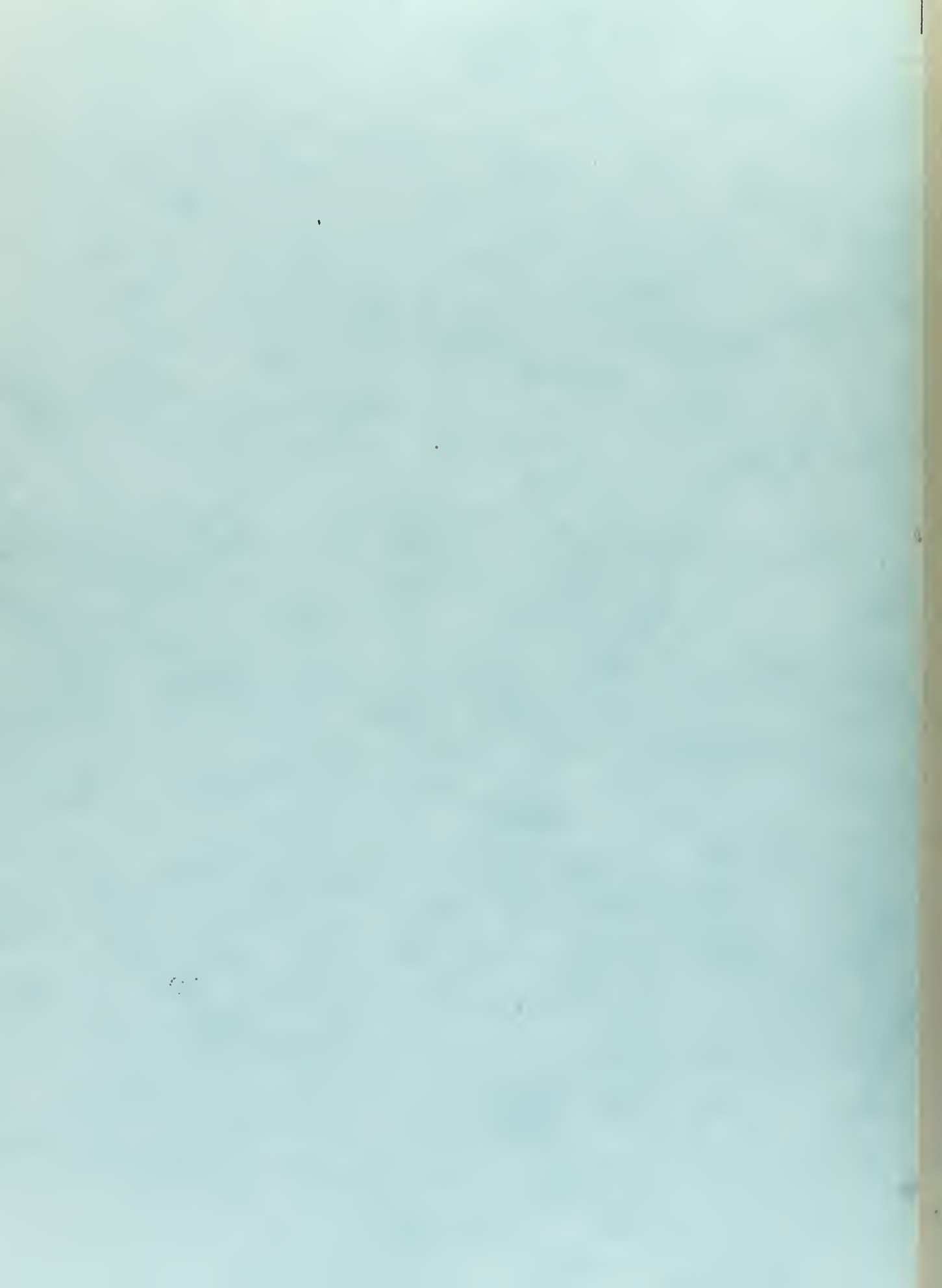


AN EXPERIMENT IN PARTITIONING SECOND-GRADE
MATHEMATICS CLASSES INTO SMALLER SECTIONS OF
COMPARABLE ABILITY

Theodore Thomas Bean



United States Naval Postgraduate School



THE SIS

AN EXPERIMENT IN PARTITIONING SECOND-GRADE
MATHEMATICS CLASSES INTO SMALLER SECTIONS OF
COMPARABLE ABILITY

by

Theodore Thomas Bean

Thesis Advisor:

R. R. Read

September 1971

1142506

Approved for public release; distribution unlimited.

An Experiment in Partitioning Second-Grade
Mathematics Classes into Smaller Sections of
Comparable Ability

by

Theodore Thomas Bean
Captain, United States Marine Corps
B.S.C.E., University of Mississippi, 1966

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1971

Thesis
B288
c. 1

ABSTRACT

The mathematics program of an elementary school was changed so that the classes were partitioned into smaller sections of comparable ability. This change reduced the pupil-teacher ratio and the length of the instruction period. The proposal produced an unanswered question: Will there be an increased achievement gain during the school year which can be attributed to the program change? Two schools served as control groups for the experiment. The California Achievement Test, 1970 Edition, was used in a pretest-posttest design to measure achievement levels before and after the elapsed time of the experiment. The question was answered in the affirmative through the use of various statistical techniques: randomized matched subjects design, analysis of covariance and the non-parametric Mann-Whitney U test. The several techniques were applied since no single standard practice was available for this problem.

TABLE OF CONTENTS

| | | |
|-------|---|----|
| I. | INTRODUCTION | 7 |
| II. | PUPIL-SCHOOL BACKGROUND | 11 |
| III. | ASSIGNMENT TO CLASSES AND INSTRUCTIONAL LEVELS | 13 |
| IV. | THE CALIFORNIA ACHIEVEMENT TEST | 16 |
| V. | METHODS OF ANALYSIS | 19 |
| | A. RANDOMIZED MATCHED SUBJECTS DESIGN | 19 |
| | B. ANALYSIS OF COVARIANCE | 22 |
| | C. MANN-WHITNEY U TEST | 24 |
| VI. | INITIAL RESULTS | 26 |
| VII. | DISCUSSION OF THE RESULTS AND FURTHER ANALYSIS | 30 |
| VIII. | CONCLUSIONS | 36 |
| | APPENDIX A: SCHOOL CHARACTERISTICS QUESTIONNAIRE | 38 |
| | APPENDIX B: SCHOOL DISTRICT MINIMUM MATH PROGRAM - GRADE 2 | 42 |
| | APPENDIX C: TEST SCORES | 44 |
| | LIST OF REFERENCES | 49 |
| | INITIAL DISTRIBUTION LIST | 51 |
| | FORM DD 1473 | 52 |

LIST OF TABLES

| | | |
|-------|---|----|
| I. | Results of School Characteristics Questionnaire --- | 11 |
| II. | Estimates of the Means and Standard Deviations ---- | 27 |
| III. | Analysis of Covariance, Total Score ----- | 28 |
| IV. | Confidence Intervals for Contrasts, Total Score --- | 29 |
| V. | Analysis of Covariance, Concept Score ----- | 30 |
| VI. | Confidence Intervals for Contrasts, Concept Score - | 31 |
| VII. | Analysis of Covariance, Computation Score ----- | 31 |
| VIII. | Confidence Intervals for Contrasts, Computation Score ----- | 31 |
| IX. | Regression Analysis on Total Score ----- | 32 |
| X. | Analysis of Covariance Total Score. School A Separated into Two Clusters ----- | 34 |
| XI. | Confidence Intervals for Contrast, Total Score, School A Separated into Two Clusters ----- | 34 |

LIST OF FIGURES

1. Scatter diagram and regression analysis; school A -- 35

ACKNOWLEDGEMENT

The author is deeply indebted to Douglas J. McRae of CTB/McGraw-Hill for his knowledge, assistance, and the loan of reference material. Specific indebtedness was incurred for his School Characteristics Questionnaire and his cluster analysis computer program which he developed as part of his doctoral dissertation while at the University of North Carolina, Chapel Hill. The author is further indebted to CTB/McGraw-Hill for supplying their California Achievement Test, 1970 Edition, and for permission to reproduce portions of the standardization data for the California Achievement Test, 1970 Edition.

I. INTRODUCTION

On September 15, 1970, the author attended a PTA meeting at his child's school. The primary purpose of the meeting was to discuss a proposal for a change in the mathematics instruction at the second grade level and to obtain the approval of the parents involved.

The proposal was primarily directed at the large pupil-teacher ratios, approximately 30 to 1, existing in the five second grade classrooms. Essentially, the pupils in the second grade would be divided into ten sections (vice five classes), with five of the sections being taught in the morning and five in the afternoon. The same five classroom teachers would be teaching both sessions. The pupil teacher ratio would be virtually halved by the proposed change. Also the smaller sections were to be composed of pupils of comparable ability. The parents were unaware of this aspect of the proposal.

A scheduling problem was inherent in the proposal due to the morning and afternoon sessions and the limited number of teaching hours available during the day. This problem was solved by having the second graders arrive and depart the school at times different from the normal school routine and by reducing the normal 40 minute math period to 30 minutes in the morning session and to 35 minutes in the afternoon session. These scheduling changes allowed the teachers to remain within their allotted teaching hours.

The main point of concern to the parents centered around the possible inconvenience caused by the different arrival and departure times of the second graders involved. In the discussion that followed it was implicitly assumed that the achievement level of the pupils would be higher under the proposed change; this assumption was never questioned. The primary reason for the assumption was the reduced pupil-teacher ratio made possible by the increased number of sections. The reduction of the math periods by ten minutes in the morning session and five minutes in the afternoon session was not felt to be significant by comparison.

The proposal was adapted and implemented the following week without the question of increased achievement being considered. The objective of this thesis, then, is to determine if there was a significant increase in mathematical ability of the second graders involved which could be attributed to the change in mathematics instruction.

Measuring and evaluating growth objectively is a classical problem in psychometrics. It has been debated extensively in the literature by Lord [Ref. 1, 2, and 3], McNemar [Ref. 4], Cronbach and Furby [Ref. 5] and others. Reference 6 is a collection of papers all dealing with the same problem - change: how to measure and evaluate it. Cronbach and Furby have even gone so far as to suggest that measurement of change should not be done. They do present a model for analyzing individual change which they consider better than

other available models, but they recommend that the results be interpreted with caution. The purpose here is neither to add nor subtract from the sometimes analytical, sometimes philosophical, arguments found in the literature but to proceed with models that have been used in the past to measure and evaluate change.

It was probable that mathematical skills would increase over the school year with or without the implementation of the proposal. This created a problem in isolating that portion of achievement, if any, due to the adopted proposal. This and other extraneous factors led to the selection of a control group. Two requirements were considered necessary for the control group. The first was that the pupils be assigned to classes in the normally accepted manner and that classes not be divided into smaller sections. The second requirement was that the pupils in the control group be of comparable ability with those in the experimental group. The first requirement seemed easy enough to satisfy, but the second seemed very difficult without prior knowledge of mathematical ability of both groups.

An objective measuring device was also needed to measure mathematical ability. The most commonly accepted way of measuring knowledge in educational institutions is by the use of tests, be they written, oral, or other. It was decided to give a written pretest and posttest to both the experimental and control groups. It was felt that this pretest-posttest design would alleviate some of the difficulties in meeting the second requirement mentioned above.

Fortunately, the author was referred to CTB/McGraw-Hill, Monterey, California, by his thesis advisor. Douglas J. McRae, of the research department of that firm, took an interest in this project. Upon his urging, the firm agreed to furnish their California Achievement Tests, 1970 Edition, to the schools involved. Consultation with Mr. McRae and the principal of the experimental school also led to the selection of a control group.

It is widely accepted that a positive correlation exists between socio-economic level and academic achievement. It was, then, on the basis of assumed socio-economic level that the experimental and control groups were deemed to be of comparable ability. A questionnaire furnished by Mr. McRae was used by the author to establish the socio-economic levels of the two groups.

The experimental group happened to be in the largest elementary school in the area, so two schools were chosen to form the control group. This allowed the numerical population of the experimental and control groups to be of the same magnitude and the results to have a broader base. The principals of the two schools tentatively selected as the control group were contacted and both were eager to cooperate in the experiment. The experimental school was labeled school A and the control schools were labeled B and C.

II. PUPIL-SCHOOL BACKGROUND

A School Characteristics Questionnaire [Appendix A], furnished by Douglas J. McRae of CTB/McGraw-Hill, was given to the principal of each school to complete. The questionnaire requested information on a variety of pupil, staff, and physical plant characteristics. The schools were all considered to be located in residential suburbs. The pertinent results of the questionnaire are presented in Table I. The number in parenthesis refers to the questionnaire item.

| SCHOOL (ENROLLMENT) | | | |
|---------------------|--------|--------|--------|
| QUESTION AREA | A(900) | B(400) | C(600) |

STUDENT DEMOGRAPHIC:

| | | | | |
|------------------|------|-----|------|------|
| Mobility | (2) | 35% | 10% | 15% |
| PTA Attendance | (4) | 20% | 20% | 10% |
| Employed Mothers | (6) | 5% | 25% | 40% |
| Per Cent White | (10) | 98% | 95% | 95% |
| Kindergarten | (13) | 75% | 100% | 100% |
| One Parent | (14) | .2% | 5% | 10% |
| English Second | (15) | .4% | 1% | 2% |
| Professional | (16) | 75% | 50% | 45% |
| White Collar | (16) | 15% | 30% | 37% |
| Skilled | (16) | 10% | 15% | 15% |
| Unskilled | (16) | 0% | 5% | 3% |

PHYSICAL PLANT AND ADMINISTRATIVE:

| | | | | |
|--------------|------|---------|---------|---------|
| Plant Age | (3) | 17 yrs | 5 yrs | 20 yrs |
| New Programs | (8) | 2-5 mos | 2-6 mos | 2-6 mos |
| Library | (17) | 7650 | - | 9000 |

STAFF CHARACTERISTICS:

| | | | | |
|----------------------|------|----------|----------|----------|
| Principal's Salary | (11) | \$17,145 | \$17,500 | \$16,500 |
| Average Start Salary | (12) | 7,500 | 7,500 | 7,500 |
| Average Salary | (19) | 9,850 | - | 10,000 |
| Average Experience | (18) | 7 years | 7½ years | 9 years |

TABLE I. RESULTS OF SCHOOL CHARACTERISTICS QUESTIONNAIRE.

It is apparent that the schools differ in several areas (e.g., mobility, employed mothers, and the number of one parent homes). There is an apparent positive correlation between the per cent of working mothers and the per cent of homes with one parent. The experimental school had a lower percentage of pupils who had attended kindergarten. This agreed with the higher mobility figures for that school, as it is plausible that a large percentage of the pupils came from a school district outside of California, where kindergarten is not required. Two of the schools (A and C) have considerably older physical plants, but all three schools maintain a clean, neat appearance conducive to learning.

In other areas the schools were very similar. They were homogeneous racially; at least 95 per cent of the pupils from each school were white. They were all members of the same school district; hence they had the same minimum math program [Appendix B]. The average experience of the teaching staff was between seven and nine years. At least 80 per cent of the pupil's parents were in white-collar or professional occupations. From this it can be assumed that the socio-economic level of the three schools was approximately the same, partially confirming the basis on which the control schools were chosen.

III. ASSIGNMENT TO CLASSES AND INSTRUCTIONAL LEVELS

The school district had set forth a statement of minimum goals and skills to be achieved by the second grade in mathematics [Appendix B]. All three schools involved in the experiment used Modern School Mathematics - Structure and Use, published by the California State Department of Education, as the basic text. This text adequately covers the minimum requirements and exceeds them, continuing into multiplication. More advanced texts are available for the gifted child, who is allowed to proceed at an accelerated pace. The advanced texts extend the concepts covered in the basic text and introduce the pupil to division.

In the experimental group, pupils were initially assigned to one of ten mathematics sections on the basis of reading performance and a departmental math test. Some assignments were changed at two different times during the school year. The first shift occurred after the pretest was given. The pupils who scored way above the average for their section were shifted to an appropriate section. No pupils were transferred to a lower section on the basis of the pretest.

The second assignment change occurred during the middle of the school year. Several pupils transferred to and from the school. Persons in authority felt that the new pupils should not be placed in advance sections; not enough information was available for basing the higher assignment.

Hence new students were placed in the lower sections. The better pupils in the lower sections were shifted to advanced sections to keep the pupil-teacher ratios comparable. It was felt that by shifting pupils in this manner no one would be in a section in which the material presented was beyond his capability. Unfortunately, the records as to which pupils were shifted are not complete. This led to a difficulty in interpreting the data, and the question is addressed later in the thesis.

Five of the sections were taught during the morning and five during the afternoon. This division of the five second grade classes into ten sections automatically reduced the pupil-teacher ratio and, as noted previously, grouped pupils of comparable ability in the same section. This sectional assignment by ability allowed the teacher to present instruction at the average achievement level of their section rather than the average level of the second grade taken as a whole.

Basically, there were five levels of instruction at the experimental school. Level I was taught at an advanced level to include multiplication and division. Levels II and III were taught at progressively lower levels. Level IV was considered to be composed of pupils who were capable of learning at the normal second grade level. This group was introduced to multiplication near the end of the school year. Level V consisted of the under-achievers. They were taught at the second grade level but with a more repetitious learning process and extensive drills in fundamentals.

In the control schools, levels were not used. Children were assigned to class more or less at random; there was no distinction made between slow and fast achievers. During the school year the level of instruction was based upon the average level of achievement within the class. The teachers gave special assistance to those children who were lagging and to those who were capable of progressing more rapidly. This assistance allowed the pupils in the control schools to have essentially the same range of instruction as that of the pupils in the experiment group.

The division of the five classrooms in the experimental school into ten sections formed the basis for the experiment. These sections allowed for five different levels of instruction and a reduced pupil-teacher ratio of 15 to 1 for levels I, II, III, and IV and 10 to 1 for level V. This is to be contrasted with the control group schools where pupils were assigned more or less at random to a class, one general level of instruction prevailed, and the pupil-teacher ratios were approximately 25 to 1 and 30 to 1 in schools B and C respectively. Further, the pupils in the experimental school were being taught five or ten minutes less per day¹ than the pupils in the control schools.

¹ Ten minutes for the morning classes and five minutes for the afternoon classes.

IV. THE CALIFORNIA ACHIEVEMENT TEST

The experimental and control groups participating in the experiment and the experiment itself have been described. Information was presented which indicates that the pupils involved were of the same socio-economic level. Presumably, then, the pupils of the different schools were, on the average, at the same level of learning at the start of their second grade year. It is recognized, however, that the maturity level of second graders can range from kindergarten to the third or fourth grade, i.e., it was not assumed that all pupils had reached the same level of academic achievement.

Considering the similarity of starting levels, then, one would expect the same magnitude of achievement in mathematical skills during the second grade school year for both groups, if there were no real difference between the math programs of the experimental and control groups.

As mentioned previously, the California Achievement Test, 1970 Edition, was the measuring device used to measure the achievement level at the beginning and end of the experiment. The test is designed for the measurement, evaluation and analysis of school achievement in the three basic skill areas: reading, mathematics, and language.

These three skill areas are measured separately. The math section of the CAT measures [Ref. 7]:

"-the ability to understand the meaning of the material presented,

- the performance of the student in applying rules, facts, concepts, conventions, and principals of problem solving in the basic curricular material, and
- the level of performance of the student in using the tools of mathematics in progressively more difficult situations."

How well does the CAT measure mathematical achievement?

This question is partially answered by Reference 8. Further information, based upon national standardization data, will be available when the Technical Report, California Achievement Test, 1970 Edition, [Ref. 9], is published. The author has had limited access to this material through Mr. McRae, and from all reports the test is well established.

Reference 7 contains tables to assist in transforming raw scores (total number of right responses) into Achievement Development Scale Scores (ADSS). The scale scores were produced by applying the Thrustones absolute scaling procedure to the standardization data [Ref. 8 and 10]. There are five levels of the CAT designed to cover grades one through twelve. The scale scores have the advantage of being articulated across all levels of the CAT, i.e., theoretically a first graders achievement can be compared with that of a student in high school. Further advantages of scale scores over raw scores are that they have interval scale properties and are normally distributed [Ref. 8]. Hence, these scores are advantageous for use in statistical analysis of the measurement of growth, trends, etc.

There are two levels of the CAT which pertain to the second grade. Level 1 is designed to test pupils in the

first and second grade. Level 2 covers the second through fourth grade, thus overlapping level 1 at the second grade. It was decided to give level 1 at the beginning of the school year and level 2 at the end of the school year. The reason two levels were chosen instead of using level 1 twice, was to allow for maximum spread of the scores over the learning period and to avoid practice effects. Interlevel articulation is a correlation measure of how well the different levels measure the same attribute (in this case, mathematical achievement) at the overlap of the levels. The interlevel articulation coefficient for levels 1 and 2, CAT, 1970 Edition, is .78 [Ref. 9].

Level 1 of the CAT was given to all participating schools during the second week of October, 1970. Level 2 was given during the third week of May, 1971, an elapsed time of seven months. The tests were given, monitored and scored by the schools in accordance with Reference 7. The raw scores were collected by the principals of the participating schools and given to the author within a reasonable period of time after each testing period.

Each level of the CAT mathematics section is divided into two parts: concepts and computational ability. The raw scores of each area were kept separate. The author converted the raw scores into scale scores. These scores for all pupils are contained in Appendix C.

V. METHODS OF ANALYSIS

The stated objective of this thesis was to measure the achievement gain in mathematical skills of three different sets of pupils, compare these measures and determine by some means whether any one of these measures is significantly different from the others. As noted in the Introduction, a standard model for answering this question has not been obtained to everyone's satisfaction.

Hence, due to a lack of a totally defensible mathematical model, several statistical techniques were applied to the data. This use of several techniques had the advantage of compensating for weaknesses in different procedures and of developing experience in the various statistical techniques. Naturally the results would be highly correlated and the reinforcing value of a second or third successful procedure is slight (one should not perform multiple tests on a single set of data). Joint probability statements cannot be made, only individual (or marginal) ones. It was still held to be valuable to have the several tests as backups for one another. The three models used in the analysis were the Randomized Matched Subject Design, the Analysis of Covariance, and the non-parametric Mann-Whitney U Test. The reader is asked to judge for himself as to which technique supports the experimental paradigm.

A. RANDOMIZED MATCHED SUBJECT DESIGN [Ref. 11 and 12]

In this test there must be an equal number of pupils in the experimental and control groups. Schools B and C were

combined to form a control group of 113 subjects. School A had 94. Random number tables were used to discard the 19 excess pupils from the control group. After the pretest was given, the pupils were ranked within the experimental and control groups on the basis of their pretest score. The two groups were then paired by placing the pupil scoring highest in the experimental group with the pupil scoring highest in the control group, etc., until 94 pairings were obtained. After the posttest was given and scored a difference score was obtained on each pair in the following manner:

$$D_i = (Y_{E_i} - Y_{C_i}) - (X_{E_i} - X_{C_i}),$$

where

X_{C_i} = pretest of i^{th} pupil in the control group

Y_{C_i} = posttest of i^{th} pupil in the control group

X_{E_i} = pretest of i^{th} pupil in the experimental group

Y_{E_i} = posttest of i^{th} pupil in the experimental group.

With a difference score thus defined, the question of whether or not the program of the experimental group was better than that of the combined control group can be addressed with the following statistical hypotheses:

$$H_0: \mu = 0$$

$$H_1: \mu > 0$$

where μ is the mean of the difference scores. If during the analysis, H_0 , the null hypothesis, is not rejected, this will

imply that no significant difference exists in the programs of the experimental and control groups. It was assumed that the difference scores had a normal distribution with a mean of zero and an unknown variance. Under these assumptions, the Student-t distribution yields the uniformly most powerful test against all one-sided alternative hypotheses. It was used in this thesis to test the null hypothesis presented above.

The randomized matched subjects design has a distinct advantage over a design where the difference between the estimated gains made by the experimental and control groups is hypothesized to be zero. The pairing on the pretest scores introduces a high correlation between the groups on posttest scores. This correlation decreases the estimate of the variance by an amount equal to twice the covariance of the means. Thus the randomized matched subjects design is more sensitive to differences between the two groups than a design where pairing is not done, since the denominator of the critical ratio is reduced.

The primary disadvantage of the randomized matched subjects design is that an implicit assumption is made, i.e., the scores obtained on the pretests and posttests are true scores. In reality the obtained score on any test can be thought of as a true score plus error, where the error can be modeled as a normally distributed random variable with zero mean and unknown variance. The difference score for each pupil, as defined, is the difference of two normally

distributed random variables. The sum of these differences has a normal distribution but the variance may change from pupil to pupil. Thus the Student-t procedure will estimate a variance which is actually a composite of each students variance rather than the value common to all. Hence this procedure for comparing growth may be questioned.

B. ANALYSIS OF COVARIANCE [Ref. 13]

The second method used in the analysis of the different mathematics program was the Analysis of Covariance. In the analysis the posttest score is considered to be dependent upon the pretest score, which is called the concomitant variable. The Analysis of Covariance takes advantage of the information furnished by the concomitant variable in a sense that the posttest scores are "corrected" by taking into account the differences between the groups on the concomitant variable. After the posttest scores are "corrected" the procedure reduces to a straight analysis of variance. Thus the results from the analysis of covariance more accurately reflect any differences that are due to the different programs than a straight analysis of variance would on the posttest scores.

The assumptions of the analysis of covariance are:

1. A random sample of size one is drawn from each of N populations.
2. Each of the populations is normal and has the same variance.

3. The population means within each group lie on a straight line and the slope of the line is the same for each group.

Since the concomitant variable is measured with error (i.e., the obtained scores are not necessarily true scores) it must also be assumed that the assumptions hold for all possible values of the concomitant variable.

The model for the analysis of covariance can be written as:

$$y_{ij} = \mu + \beta_j + \gamma (x_{ij} - \bar{x}_{..}) + e_{ij} \quad \begin{matrix} i=1,2,\dots,n_j \\ j=1,2,\dots,r \end{matrix}$$

$$e_{ij} \sim N(0, \sigma^2)$$

$$\sum_{j=1}^r n_j \beta_j = 0 \quad N = \sum_{j=1}^r n_j$$

where

r = the number of schools.

n_j = the number of pupils in the j^{th} school.

x_{ij} = pretest of i^{th} pupil in the j^{th} school.

y_{ij} = posttest of i^{th} pupil in the j^{th} school.

β_j = program effect of j^{th} school.

γ = slope of regression line (posttest on pretest).

μ = the average of all r population means.

The objective of the analysis is to choose between the following set of hypotheses:

$$H_0: \beta_j = \beta \quad j=1,2,\dots,r.$$

$$H_1: \text{not all } \beta_j \text{ are the same}$$

The null hypothesis implies that the corrected program effects are all equal.

The robustness of this test is not as well known as that of the straight analysis of variance. Feldt [Ref.14], in a paper comparing several techniques, stated that the number of assumptions required for a valid analysis of covariance renders the technique generally less applicable than other statistical tests used for the same purpose. The failure of the data to meet the assumptions is thought to be more serious in analysis of covariance than in straight analysis of variance (the Student-t in the two sample case) especially with failure to satisfy regression assumptions.

Feldt investigated the precision of a factorial design and of the analysis of covariance. He discovered that for regression correlations of greater than 0.6 analysis of covariance has the greater precision. Since CTB/McGraw-Hill has estimated this correlation (their interlevel of articulation) to be 0.78^2 in their standardization procedures, it was believed that the analysis of covariance in this thesis would yield valid results.

C. MANN-WHITNEY U TEST

A third test was presented for those readers unwilling to accept the assumptions in the analysis of covariance or the validity of the randomized matched subjects design. It

² See the r values in Tables IX and Figure 1 for values obtained during the analysis.

is the non-parametric Mann-Whitney U test. In this test it was again necessary to form the two control schools into one, but it was not necessary to delete the 19 pupils as in the randomized matched subjects design. The achievement gain score (posttest-pretest) was computed for all pupils. Again the thesis in question could be answered by a set of hypotheses

H_0 : The achievement gains for the experimental and control groups have the same distribution.

H_1 : The achievement gains for the experimental group are stochastically larger than those of the control group.

There are no assumptions made as to the distribution of the achievement gains other than the distributions are continuous and that the observations are completely independent. Siegel [Ref. 15] states that the power efficiency of the Mann-Whitney U test approaches 95.5 per cent for large sample sizes compared to the Student-t. It is, therefore, an excellent alternative to the Student-t, and it does not have the restrictive assumptions associated with the analysis of covariance.

VI. INITIAL RESULTS

The set of hypotheses presented in each of the three different techniques of analysis were all designed to answer the same question: Was the achievement level in mathematics of the pupils in school A larger than the gain normally attributed to the learning process during the second grade? If the null hypothesis was not accepted during the analysis it would be a strong indication that the change in mathematical programs at school A contributed significantly to the gains made by the pupils. If the opposite occurred, it could be concluded that the program had no significant effect.

Achievement Development Scale Scores were used in all statistical analysis and the significance level of all tests was arbitrarily set at the 5 per cent level. This significance level is the probability of a Type I error, i.e., the probability that the null hypothesis is rejected when in reality it is true. The power of the tests, i.e., the probability of failing to reject the null hypothesis when the alternative is true, were not determined.

All three statistical techniques discussed were used on the total test score and each of the subtest scores, i.e., grasp of concepts and computational ability. The basic results are presented in Table II. The table includes the estimates of the mean and standard deviations for each of the three scores. The remaining results in this section apply only to the total scores.

| SCHOOL | COMPLETE | | | | CONCEPTS | | | | COMPUTATIONS | | | |
|--------|-----------|-------|-----------|-------|-----------|-------|-----------|-------|--------------|-------|-----------|-------|
| | PRE | | POST | | PRE | | POST | | PRE | | POST | |
| | \bar{X} | S_x | \bar{Y} | S_y | \bar{X} | S_x | \bar{Y} | S_y | \bar{X} | S_x | \bar{Y} | S_y |
| A | 280.5 | 24.0 | 342.8 | 50.8 | 327. | 38.4 | 391.2 | 49.8 | 276.9 | 20.6 | 331.2 | 46.2 |
| B | 276.4 | 27.8 | 313.3 | 39.6 | 320.6 | 37.3 | 366.2 | 46.4 | 273.6 | 25.2 | 303.7 | 35.2 |
| C | 279.1 | 30.4 | 317.8 | 35.1 | 327.7 | 38.0 | 362.3 | 43.2 | 279.8 | 28.3 | 311.6 | 31.4 |

TABLE II. ESTIMATES OF THE MEANS AND STANDARD DEVIATIONS.

In the randomized matched subject design, the estimate of the mean of 94 difference scores was 24.4. The sample standard deviation was 48.7. These figures combined to yield $t = 4.86 > 1.99$. Hence the null hypothesis was rejected.

The results of the analysis of covariance are contained in Table III. $F = 13.03 > 3.00$, hence the null hypothesis was rejected. The rejection of the null hypothesis in the analysis of covariance only indicates that all

| SOURCE | SS _x | SP _{xy} | SS _y | SS' _y | d.f. | MS' _y |
|-----------|-----------------|------------------|-----------------|------------------|------|------------------|
| TREATMENT | 528.4 | 3,814.1 | 37,606.9 | 30,165.2 | 2 | 15,002.6 |
| ERROR | 150,788.5 | 157,041.1 | 398,519.6 | 234,966.5 | 203 | 1,157.5 |
| TOTAL | 151,316.9 | 160,855.2 | 436,126.44 | 265,131.7 | | |

TABLE III. ANALYSIS OF COVARIANCE, TOTAL SCORE.

programs were not the same. A technique exists to determine which program or programs were significantly different. A linear combination of the form

$$L = c_1\beta_A + c_2\beta_B + c_3\beta_C, \quad \sum_{j=1}^3 c_j = 0$$

where the c_j 's are constants, is called a contrast. Scheffé's method [Ref. 15] was used to obtain confidence intervals for several contrasts and to determine which contrasts were significantly different from zero. β_i , $i = A, B, C$ are the program effects of the different schools. The results are contained in Table IV.

| CONTRAST | INTERVAL |
|--|--------------------------|
| $\beta_A - \beta_B$ | $10.5 \leq L \leq 39.9$ |
| $\beta_A - \beta_C$ | $10.0 \leq L \leq 37.0$ |
| $\beta_B - \beta_C$ | $-17.5 \leq L \leq 14.2$ |
| $\beta_A - \frac{1}{2}(\beta_B + \beta_C)$ | $11.8 \leq L \leq 35.2$ |

TABLE IV. CONFIDENCE INTERVAL FOR CONTRASTS, TOTAL SCORE.

It is obvious from Table IV that school A is the reason the null hypothesis was not accepted. The confidence intervals associated with school A do not contain zero and are greater than zero.

The Mann-Whitney U test yielded the same result, i.e., the null hypothesis was rejected. The achievement gain scores yielded a U statistic of 6884.5 (correction for ties was included). This transformed into a normalized variate, $z = 3.667$. The probability of this occurring is $0.0013 < 0.05$.

VII. DISCUSSION OF THE RESULTS AND FURTHER ANALYSIS

All three methods of analysis soundly rejected the null hypothesis. Hence, from the preliminary results, it was concluded that the division of classes into small sections, composed of pupils of comparable ability, significantly increased the achievement gains during the second grade school year.

At this point the author decided to determine if the programs were significantly different in both areas of the California Achievement Test. The analysis of covariance was chosen as the technique for further analysis based on the author's belief that it yielded the most information and that it was broader based since the control schools were not combined. The results of the analysis of the concept scores are presented in Table V and computation scores in Table VII. The confidence intervals for contrasts are contained in Tables VI and VIII, respectively.

| SOURCE | SS_x | SP_{xy} | SS_y | SS'_y | d.f. | MS'_y |
|-----------|-----------|-----------|-----------|-----------|------|----------|
| TREATMENT | 1,728.4 | 2,708.5 | 38,634.4 | 35,482.1 | 2 | 17,741.0 |
| ERROR | 298,927.5 | 230,133.3 | 457,982.5 | 280,811.4 | 203 | 1,383.3 |
| TOTAL | 300,655.9 | 232,891.8 | 496,616.9 | 316,293.5 | | |

TABLE V. ANALYSIS OF COVARIANCE, CONCEPT SCORE.

For concept scores the statistic $F = 12.83 > 3.00$, hence the null hypothesis was also rejected for concepts, and from Table VI it is clear that school A was the reason.

| CONTRAST | INTERVAL |
|--|-------------------------|
| $\beta_A - \beta_B$ | $3.8 \leq L \leq 36.1$ |
| $\beta_A - \beta_C$ | $14.7 \leq L \leq 44.2$ |
| $\beta_B - \beta_C$ | $-7.9 \leq L \leq 26.7$ |
| $\beta_A - \frac{1}{2}(\beta_B + \beta_C)$ | $12.0 \leq L \leq 37.5$ |

TABLE VI. CONFIDENCE INTERVALS FOR CONTRASTS, CONCEPT SCORE.

| SOURCE | SS _x | SP _{xy} | SS _y | SS' _y | d.f. | MS' _y |
|-----------|-----------------|------------------|-----------------|------------------|------|------------------|
| TREATMENT | 392.6 | 3,363.3 | 29,037.6 | 23,948.9 | 2 | 11,974.4 |
| ERROR | 122,137.1 | 95,356.0 | 320,972.5 | 250,525.3 | 203 | 1,239.1 |
| TOTAL | 122,529.7 | 98,719.3 | 359,010.1 | 274,479.2 | | |

TABLE VII. ANALYSIS OF COVARIANCE, COMPUTATION SCORE.

| CONTRAST | INTERVAL |
|--|-------------------------|
| $\beta_A - \beta_B$ | $9.8 \leq L \leq 40.1$ |
| $\beta_A - \beta_B$ | $4.0 \leq L \leq 32.0$ |
| $\beta_B - \beta_C$ | $-23.3 \leq L \leq 9.4$ |
| $\beta_A - \frac{1}{2}(\beta_B + \beta_C)$ | $9.4 \leq L \leq 33.6$ |

TABLE VIII. CONFIDENCE INTERVALS FOR CONTRAST, COMPUTATION SCORE.

From the analysis presented, thus far, it is clear that the program at school A allowed for sizeable achievement gains in all math areas of the CAT. A regression of the posttest scores on the pretest scores, total scores, was done to see how well the analysis of covariance assumptions were met. The results are contained in Table IX. The

slopes of the regression equations were within reason for all of the schools, but the correlation was low for school A³. Histograms were made for all three schools on the pre-test and posttest total scores. All appeared to be normally distributed except the posttest scores of school A. A bi-modal distribution appeared, and further analysis was suggested.

| SCHOOL | REGRESSION EQUATION | CORRELATION |
|--------|-----------------------|-------------|
| A | $Y = 18.27 + 1.16 X$ | 0.545 |
| B | $Y = - 4.40 + 1.15 X$ | 0.808 |
| C | $Y = 75.93 + 0.87 X$ | 0.750 |

TABLE IX. REGRESSION ANALYSIS ON TOTAL SCORES.

An investigation was begun to see if the apparent bi-modal distribution of posttest scores of school A was associated with the section assignments. At this time it was discovered that complete details of final section assignments were not available. The final assignments were different from the original because of shifts during the school year. The reason for the shifts were explained in detail in Section III of this thesis. The assistant principal and the only available teacher concerned with the experiment assisted

³It is possible, through statistical techniques, to test the slopes for equality and to obtain confidence intervals for the correlation. Equality of slopes was assumed in the Analysis of Covariance and the tests were not done to avoid debate about the significance level of the Analysis of Covariance.

the author in verifying 79 (out of 94) of the final assignments. This information was not considered to be complete enough for further analysis.

A cluster analysis computer program was obtained from Mr. McRae [Ref. 16] to separate the pupils into clusters based on their test scores. The program contained several different methods to accomplish this. Two of the methods were used with identical results. The program separated the pupils of school A into two clusters; basically, the clusters consisted of levels I and II, labeled cluster A; and levels III, IV, and V, labeled cluster B. Since 79 of the final assignments were known, it was easy enough to verify the soundness of the routines. Four errors were detected. Three pupils who were known to be in cluster A were assigned to B and one who was known to be in cluster B was assigned to A. The author considered the number of known errors to be sufficiently small to avoid appreciable discrepancies in further analysis.

Another Analysis of Covariance was run on the total scores. This time four groups were included in the analysis: School A-cluster A, school A-cluster B, school B, and school C. The results are presented in Tables X and XI.

The statistic $F = 99.29 > 2.6$. Hence the null hypothesis was rejected. It is clear from Table XI that the scores from school A-cluster A accounted for the rejection of the null

hypothesis. Cluster A consisted of these pupils assigned to levels I and II where the level of instruction was the most advanced.

| SOURCE | SS_x | SP_{xy} | SS_y | SS'_y | d.f. | MS'_y |
|-----------|-----------|-----------|-----------|-----------|------|----------|
| TREATMENT | 10,539.9 | 48,806.3 | 239,804.1 | 157,992.7 | 3 | 52,669.2 |
| ERROR | 140,777.0 | 112,048.9 | 196,322.3 | 107,139.0 | 202 | 530.4 |
| TOTAL | 151,316.9 | 160,855.2 | 436,126.4 | 265,131.7 | | |

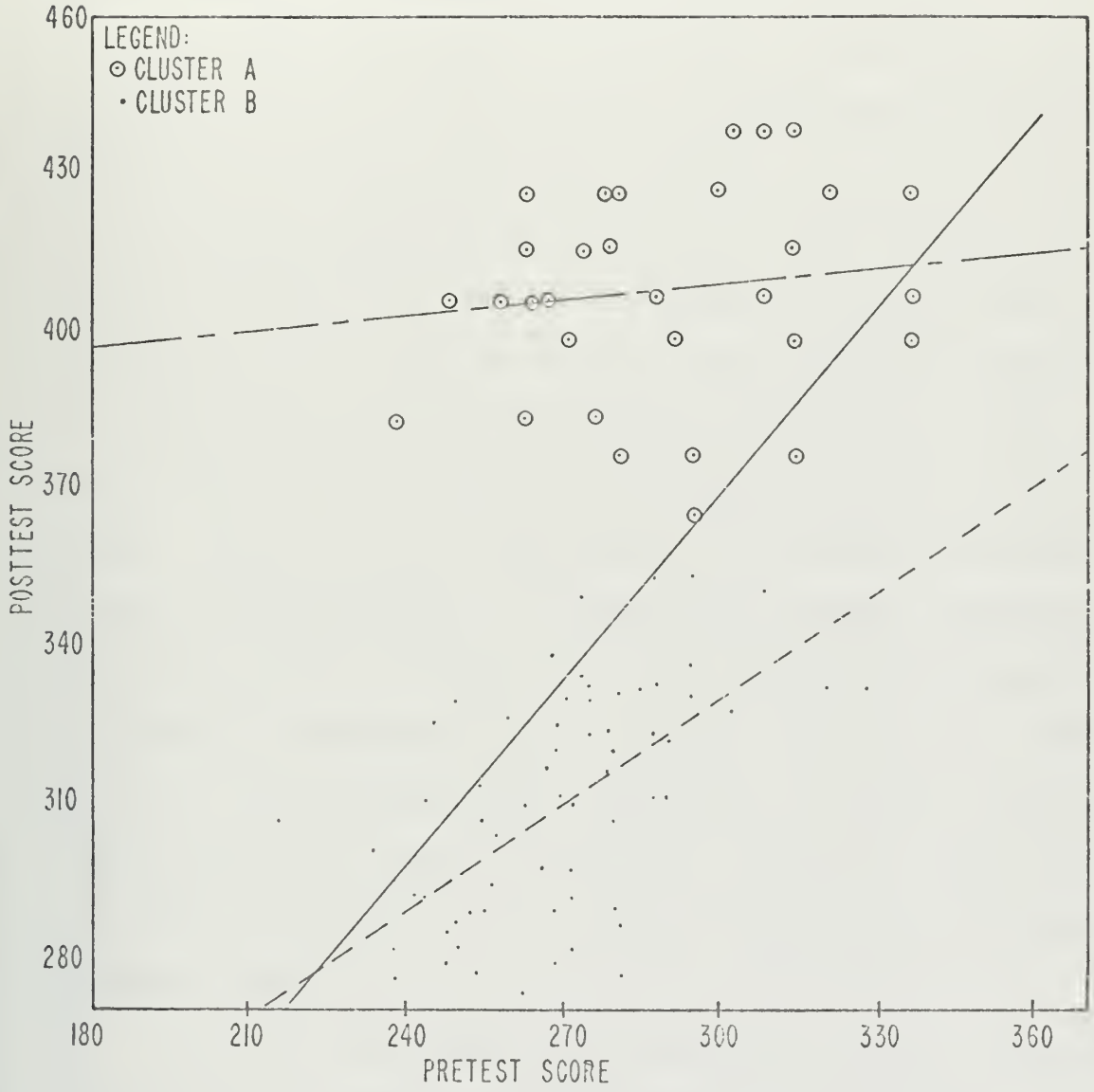
TABLE X. ANALYSIS OF COVARIANCE, TOTAL SCORE. SCHOOL A SEPARATED INTO TWO CLUSTERS.

| CONTRAST | INTERVAL |
|-------------------------|--------------------------|
| $\beta_{AB}-\beta_B$ | $-13.5 \leq L \leq 11.0$ |
| $\beta_{AB}-\beta_C$ | $-15.0 \leq L \leq 7.9$ |
| $\beta_{AA}-\beta_B$ | $64.4 \leq L \leq 94.3$ |
| $\beta_{AA}-\beta_C$ | $62.8 \leq L \leq 91.1$ |
| $\beta_{AA}-\beta_{AB}$ | $66.0 \leq L \leq 95.0$ |
| $\beta_B-\beta_C$ | $-14.5 \leq L \leq 9.9$ |

TABLE XI. CONFIDENCE INTERVALS FOR CONTRAST, TOTAL SCORES. SCHOOL A SEPARATED INTO TWO CLUSTERS.

The scatter diagram for school A is shown in Figure 1. The regression equations and correlations for the school as a whole and separated into clusters are shown on the figure. The correlation of the pupils in cluster B increased to 0.823, while the correlation of the pupils in cluster A decreased to 0.124.

FIGURE 1. SCATTER DIAGRAM AND REGRESSION ANALYSIS, SCHOOL A.



SCHOOL A ———
 $Y = 18.27 + 1.15 X$
 $S_Y = 50.84$
 $S_{Y|X} = 42.62$
 $r = 0.545$

CLUSTER A - - - - -
 $Y = 378.8 + 0.097 X$
 $S_Y = 19.43$
 $S_{Y|X} = 19.28$
 $r = 0.124$

CLUSTER B - . - . -
 $Y = 129.48 + 0.659 X$
 $S_Y = 30.43$
 $S_{Y|X} = 17.29$
 $r = 0.823$

VIII. CONCLUSIONS

In the Introduction it was questioned whether a proposed change in a mathematics program would lead to increased math achievement. The author contends that through statistical analysis this thesis established the increased achievement beyond a reasonable doubt.

From the confidence intervals in Table XI, it is clear that no significant differences existed between cluster B of school A and schools B and C; the confidence intervals for the appropriate contrasts all contain zero. This result, along with the information that cluster B at school A had a sample standard deviation of 30.4 on the posttest score, led the author to conclude that no one was hurt by the program. It afforded the slow achievers the opportunity to receive concentrated instruction at their level. The fact that the concentrated instruction at level V assisted those pupils to advance faster was reflected by the low variability of cluster B's posttest scores compared with schools B and C.

As far as the author can ascertain from the analysis, the average student was unaffected by the program. It is readily apparent that those pupils in levels I and II received the greatest benefit. The means for levels I and II, on the posttest score, was above the 99th percentile of the nation based upon the standardization data of the CAT. On the average, cluster A missed less than two questions

per child out of 117 on the posttest. This spectacular result further suggests that cluster A could have scored even higher on an expanded test. Since the CAT was designed to be used through the fourth grade, the second graders are to be commended on their accomplishment.

It may be questioned whether or not the teachers at the experimental school were concentrating their instruction in the area of the CAT. The author talked to one of the teachers in the experimental school and to the principals of all three schools. He was assured that "teaching to the test" was not done at any of the schools.

It is noted, that, since the experimental teaching program was conceived and implemented by school A teachers, the so called "Hawthorne effect" [Ref. 17] may have been operative. The results are quite striking, however, and the author feels that if such a program were administered by other teachers the improvement indicated here would not be diminished by any large degree. Finally, it is noted that none of the students at any of the schools were aware of the experiment; only the teachers.

It is hoped that this thesis is of some aid in keeping the program at school A; including the extension of it into the other elementary grades. It is further hoped that it can be used as a basis to implement the experimental program at other schools with the school district.

APPENDIX A: SCHOOL CHARACTERISTICS QUESTIONNAIRE

1. How many students are enrolled in this school at each of the following grade levels:

| | |
|---|----|
| K | 7 |
| 1 | 8 |
| 2 | 9 |
| 3 | 10 |
| 4 | 11 |
| 5 | 12 |
| 6 | |

2. About what percentage of the students who attended this school last year are no longer attending this school (do not count those who have moved because of graduation or are being bussed to other schools)?

_____ %

3. How old is the main classroom building of this school plant?

_____ years old.

4. About what percent of the families of students at this school are represented at a typical meeting of the PTA or similar parent group?

_____ %

5. Which of the following categories best describe the neighborhood served by this school?

- _____ a. rural
- _____ b. residential suburb
- _____ c. industrial suburb
- _____ d. small town (5,000 or less)
- _____ e. city of 5,000 to 50,000
- _____ f. residential area of a large city (50,000+)
- _____ g. inner part of a large city (50,000+)

6. About what percentage of students in this school have mothers who are employed outside of the home?

_____ %

7. From which of the following groups (check all that apply) is formal approval required to initiate new education programs in this school (e.g., team teaching, new curricula, ungraded classrooms, resource rooms, etc.)?

- _____ Board of Education
- _____ Superintendent
- _____ District administration
- _____ other than Superintendent
- _____ Parents
- _____ Teachers
- _____ No formal approval needed

8. About how long does it usually take to implement a new educational program in this school (i.e., from the time the decision is made to adopt it until the time it is actually introduced)?

_____ months

9. (a) (Elementary Schools) What is the copyright date of the regular class reading book used in the third grade at this school? _____

(b) (Junior and senior high schools) What is the copyright date of the regular American history text used in this school? _____

10. About what percentage of the students in this school are white?

_____ %

11. What is the annual salary of the principal of this school?

\$ _____

12. What is the starting annual salary of a fully certified beginning teacher in this school system?

\$ _____

13. (Elementary schools only) About what percentage of the students now in Grade 1 in this school attended Kindergarten or its equivalent?

_____ %

14. About what percentage of the students in this school are living in homes in which there is only one parent?

_____ %

15. About what percentage of the students in this school speak a language other than English outside of school or come from homes in which a language other than English is spoken most of the time?

_____ %

16. About what percentage of the pupils served by this school fall into each of the categories listed in the chart below (the total should equal 100%)?

| Occupational Category | % |
|--|------|
| children of professionals and managers (doctors, lawyers, engineers, executives, etc.) | |
| children of white collar workers other than those in (a) above (proprietors, salesman, clerks, etc.) | |
| children of skilled workers (electricians, carpenters, repair men, factory workers, etc) | |
| children of unskilled workers (laborers, janitors, dishwashers, etc.) | |
| TOTAL | 100% |

17. About how many catalogued volumes are there in the library of this school?

_____ volumes

18. What is the average full-time teaching experience of the teaching staff of this school (consider counseling as teaching experience)?

_____ years

19. What is the approximate average annual salary of the teaching staff in this school?

\$ _____

20. Please estimate in the chart below the number of hours per week that each of the specified kinds of people are working in this school:

| Type of person | # hours per week |
|--|------------------|
| Guidance Counselor | |
| Psychologist | |
| Child Welfare and Attendance Office | |
| Nurse | |
| Speech Therapist | |
| Remedial Reading Specialist | |
| English-Second- Language Specialist | |

hours per week

Art Teacher
Sex Education Consultant
Librarian
Teacher Aides

APPENDIX B: SCHOOL DISTRICT MINIMUM MATH PROGRAM - GRADE 2

Students in the second grade

- will be able to read numerals through 200 and write numerals through 100, and to count to 100 by 1's and 10's, and to 20 by 5's.
- will gain a working knowledge of the concepts of numbers 0 through 99 and be able to identify the number before and after a given number through 99. They will have some experiences with numbers 100 through 999 including the identification of place value through the hundreds place.
- will be exposed to number words through "ten."
- will gain a working knowledge of the ordinal numerals first through seventh (using teacher-made materials).
- will be able to demonstrate an understanding that there are many ways to write the numeral for a given number.
- will master the addition and subtraction combinations through 15, with emphasis on vertical notation, and will be able to complete the addition and subtraction sentences within these number families.
- will be able to demonstrate an understanding of and be able to add and subtract two-digit numbers which do not require carrying or borrowing.
- will be exposed to and have some practice with addition of three-digit numbers with no carrying.
- will be exposed to and have some practice with addition problems involving carrying and with subtraction problems involving borrowing.
- will be able to demonstrate an understanding of, and be able to work addition problems with, three addends up to and including two-digit numbers.
- will be able to demonstrate an understanding of the meaning of the mathematical vocabulary and symbols included in the minimum program for grade two.
- will be able to demonstrate an understanding of and be able to work oral story problems using the concepts and skills mastered.

- will be able to demonstrate an understanding of the concepts of $\frac{1}{2}$ and $\frac{1}{4}$ and be exposed to $\frac{1}{3}$ as part of a whole.
- will be able to tell time to the nearest hour and half-hour.
- will be able to use pennies, nickels, dimes, quarters, and dollars.
- will have further exposure to a calendar and will develop further understanding of a day, week and month in measuring time.

APPENDIX C: TEST SCORES

SCHOOL A

| STUDENT NUMBER | PRETEST | | | POSTTEST | | |
|-------------------|----------|-------------|-------|----------|-------------|-------|
| | CONCEPTS | COMPUTATION | TOTAL | CONCEPTS | COMPUTATION | TOTAL |
| 1 | 322 | 303 | 290 | 333 | 313 | 310 |
| 2 | 378 | 293 | 316 | 444 | 373 | 397 |
| 3 | 346 | 258 | 271 | 333 | 282 | 288 |
| 4 | 327 | 235 | 250 | 428 | 400 | 405 |
| 5 | 317 | 260 | 265 | 464 | 400 | 426 |
| 6 | 304 | 317 | 282 | 464 | 400 | 426 |
| 7 | 322 | 276 | 280 | 402 | 307 | 322 |
| 8 | 339 | 280 | 290 | 393 | 309 | 322 |
| 9 | 395 | 317 | 339 | 428 | 385 | 397 |
| 10 | 346 | 273 | 287 | 428 | 313 | 331 |
| 11 | 339 | 286 | 293 | 428 | 385 | 397 |
| 12 | 395 | 317 | 339 | 464 | 373 | 405 |
| 13 | 365 | 293 | 310 | 402 | 334 | 348 |
| 14 | 339 | 293 | 297 | 444 | 348 | 375 |
| 15 | 309 | 286 | 278 | 384 | 322 | 331 |
| 16 | 365 | 286 | 305 | 402 | 313 | 327 |
| 17 | 354 | 263 | 278 | 384 | 311 | 322 |
| 18 | 292 | 268 | 264 | 444 | 400 | 415 |
| 19 | 395 | 293 | 323 | 488 | 385 | 426 |
| 20 | 365 | 293 | 310 | 488 | 400 | 437 |
| 21 | 346 | 286 | 297 | 384 | 319 | 329 |
| 22 | 339 | 268 | 280 | 364 | 307 | 315 |
| 23 | 365 | 293 | 310 | 428 | 400 | 405 |
| 24 | 354 | 317 | 316 | 402 | 373 | 375 |
| 25 | 354 | 317 | 316 | 464 | 385 | 415 |
| 26 | 263 | 254 | 244 | 328 | 288 | 291 |
| 27 | 339 | 303 | 301 | 464 | 400 | 426 |
| 28 | 300 | 249 | 252 | 276 | 299 | 281 |
| 29 | 339 | 235 | 252 | 338 | 278 | 286 |
| 30 | 296 | 258 | 257 | 348 | 299 | 305 |
| 31 | 296 | 268 | 265 | 353 | 256 | 273 |
| 32 | 309 | 280 | 276 | 377 | 326 | 333 |
| 33 | 365 | 260 | 276 | 344 | 326 | 348 |
| 34 | 283 | 263 | 257 | 343 | 311 | 312 |
| 35 | 395 | 317 | 339 | 464 | 400 | 426 |
| 36 | 332 | 303 | 297 | 402 | 338 | 352 |
| 37 | 275 | 251 | 246 | 393 | 267 | 291 |
| 38 | 304 | 280 | 274 | 348 | 286 | 296 |
| 39 | 327 | 276 | 282 | 358 | 297 | 306 |
| 40 | 309 | 303 | 282 | 414 | 301 | 319 |
| 41 | 263 | 249 | 240 | 353 | 258 | 275 |
| 42 | 296 | 280 | 271 | 358 | 303 | 310 |
| 43 | 292 | 258 | 256 | 324 | 288 | 289 |
| 44 | 292 | 303 | 274 | 319 | 280 | 281 |

| | | | | | | |
|----|-----|-----|-----|-----|-----|-----|
| 45 | 322 | 263 | 269 | 428 | 400 | 405 |
| 46 | 327 | 280 | 284 | 370 | 271 | 289 |
| 47 | 365 | 286 | 305 | 488 | 400 | 437 |
| 48 | 378 | 293 | 316 | 488 | 400 | 437 |
| 49 | 317 | 280 | 280 | 444 | 400 | 415 |
| 50 | 292 | 268 | 264 | 402 | 385 | 382 |
| 51 | 327 | 273 | 280 | 464 | 400 | 426 |
| 52 | 365 | 280 | 301 | 464 | 400 | 426 |
| 53 | 296 | 276 | 269 | 370 | 334 | 337 |
| 54 | 313 | 280 | 278 | 393 | 317 | 329 |
| 55 | 296 | 268 | 265 | 428 | 400 | 405 |
| 56 | 346 | 260 | 273 | 464 | 363 | 397 |
| 57 | 339 | 266 | 278 | 428 | 363 | 382 |
| 58 | 327 | 276 | 282 | 377 | 322 | 329 |
| 59 | 317 | 273 | 276 | 444 | 400 | 415 |
| 60 | 300 | 260 | 260 | 428 | 400 | 405 |
| 61 | 339 | 280 | 290 | 428 | 400 | 405 |
| 62 | 279 | 268 | 260 | 338 | 301 | 303 |
| 63 | 328 | 270 | 246 | 343 | 307 | 309 |
| 64 | 300 | 280 | 273 | 428 | 311 | 329 |
| 65 | 263 | 249 | 240 | 310 | 284 | 281 |
| 66 | 271 | 270 | 259 | 314 | 299 | 294 |
| 67 | 271 | 260 | 251 | 319 | 284 | 284 |
| 68 | 300 | 276 | 271 | 364 | 313 | 319 |
| 69 | 296 | 293 | 274 | 377 | 292 | 308 |
| 70 | 322 | 263 | 269 | 384 | 301 | 315 |
| 71 | 292 | 286 | 271 | 377 | 315 | 324 |
| 72 | 271 | 239 | 235 | 343 | 292 | 329 |
| 73 | 296 | 268 | 265 | 358 | 299 | 308 |
| 74 | 292 | 249 | 250 | 348 | 264 | 278 |
| 75 | 304 | 253 | 256 | 281 | 290 | 276 |
| 76 | 327 | 251 | 261 | 377 | 317 | 325 |
| 77 | 327 | 303 | 293 | 358 | 317 | 321 |
| 78 | 339 | 258 | 269 | 348 | 288 | 297 |
| 79 | 339 | 280 | 290 | 370 | 363 | 352 |
| 80 | 365 | 233 | 255 | 370 | 296 | 288 |
| 81 | 332 | 303 | 297 | 414 | 319 | 335 |
| 82 | 395 | 303 | 330 | 402 | 317 | 331 |
| 83 | 395 | 261 | 282 | 393 | 385 | 375 |
| 84 | 365 | 317 | 323 | 414 | 315 | 331 |
| 85 | 317 | 270 | 274 | 296 | 303 | 291 |
| 86 | 322 | 286 | 284 | 353 | 260 | 276 |
| 87 | 346 | 286 | 297 | 393 | 363 | 364 |
| 88 | 292 | 248 | 248 | 393 | 311 | 324 |
| 89 | 322 | 261 | 268 | 348 | 299 | 305 |
| 90 | 332 | 286 | 290 | 377 | 324 | 331 |
| 91 | 322 | 265 | 271 | 353 | 262 | 278 |
| 92 | 395 | 317 | 339 | 428 | 400 | 405 |
| 93 | 395 | 268 | 293 | 343 | 309 | 310 |
| 94 | 322 | 286 | 284 | 364 | 269 | 286 |

SCHOOL B

| | | | | | | |
|----|-----|-----|-----|-----|-----|-----|
| 1 | 365 | 317 | 323 | 414 | 385 | 389 |
| 2 | 339 | 293 | 297 | 393 | 305 | 319 |
| 3 | 296 | 258 | 257 | 338 | 251 | 265 |
| 4 | 162 | 233 | 193 | 328 | 262 | 270 |
| 5 | 378 | 317 | 330 | 464 | 354 | 389 |
| 6 | 365 | 317 | 323 | 414 | 373 | 382 |
| 7 | 346 | 303 | 305 | 377 | 338 | 342 |
| 8 | 395 | 303 | 330 | 464 | 343 | 376 |
| 9 | 354 | 303 | 310 | 377 | 354 | 352 |
| 10 | 365 | 303 | 316 | 464 | 328 | 355 |
| 11 | 317 | 276 | 278 | 364 | 315 | 321 |
| 12 | 339 | 303 | 301 | 377 | 286 | 303 |
| 13 | 332 | 317 | 301 | 488 | 348 | 389 |
| 14 | 354 | 317 | 316 | 428 | 385 | 405 |
| 15 | 339 | 286 | 293 | 393 | 348 | 355 |
| 16 | 346 | 286 | 297 | 370 | 315 | 322 |
| 17 | 300 | 270 | 268 | 353 | 275 | 288 |
| 18 | 327 | 293 | 290 | 414 | 324 | 339 |
| 19 | 254 | 265 | 250 | 310 | 290 | 286 |
| 20 | 292 | 253 | 252 | 300 | 286 | 280 |
| 21 | 339 | 235 | 252 | 348 | 311 | 313 |
| 22 | 332 | 266 | 276 | 428 | 354 | 375 |
| 23 | 346 | 286 | 297 | 414 | 334 | 352 |
| 24 | 287 | 242 | 243 | 328 | 292 | 294 |
| 25 | 309 | 246 | 252 | 296 | 262 | 258 |
| 26 | 309 | 273 | 273 | 370 | 284 | 300 |
| 27 | 278 | 237 | 259 | 402 | 292 | 312 |
| 28 | 313 | 261 | 265 | 348 | 269 | 281 |
| 29 | 296 | 254 | 255 | 338 | 305 | 306 |
| 30 | 279 | 260 | 254 | 319 | 277 | 278 |
| 31 | 300 | 251 | 254 | 343 | 307 | 309 |
| 32 | 339 | 268 | 280 | 338 | 290 | 296 |
| 33 | 322 | 263 | 269 | 393 | 267 | 291 |
| 34 | 300 | 261 | 261 | 353 | 269 | 283 |
| 35 | 296 | 233 | 239 | 291 | 269 | 262 |
| 36 | 339 | 317 | 305 | 384 | 324 | 333 |
| 37 | 317 | 293 | 284 | 353 | 315 | 318 |
| 38 | 271 | 260 | 251 | 300 | 251 | 252 |
| 39 | 292 | 263 | 260 | 358 | 301 | 309 |
| 40 | 354 | 276 | 293 | 358 | 309 | 315 |
| 41 | 304 | 265 | 265 | 343 | 303 | 306 |
| 42 | 327 | 273 | 280 | 364 | 305 | 313 |
| 43 | 327 | 276 | 282 | 314 | 277 | 276 |
| 44 | 300 | 228 | 237 | 328 | 264 | 271 |
| 45 | 339 | 276 | 287 | 328 | 299 | 299 |
| 46 | 327 | 260 | 268 | 364 | 311 | 318 |
| 47 | 296 | 249 | 251 | 343 | 260 | 273 |
| 48 | 296 | 256 | 256 | 228 | 256 | 265 |
| 49 | 309 | 265 | 267 | 338 | 258 | 270 |

SCHOOL C

| | | | | | | |
|----|-----|-----|-----|-----|-----|-----|
| 1 | 292 | 242 | 244 | 343 | 278 | 288 |
| 2 | 378 | 251 | 269 | 384 | 313 | 324 |
| 3 | 365 | 254 | 271 | 393 | 313 | 325 |
| 4 | 296 | 270 | 267 | 338 | 262 | 273 |
| 5 | 309 | 235 | 274 | 319 | 313 | 306 |
| 6 | 339 | 286 | 293 | 338 | 264 | 275 |
| 7 | 279 | 246 | 243 | 324 | 275 | 278 |
| 8 | 313 | 254 | 260 | 370 | 311 | 319 |
| 9 | 378 | 303 | 323 | 377 | 331 | 337 |
| 10 | 304 | 246 | 251 | 333 | 284 | 289 |
| 11 | 332 | 293 | 293 | 358 | 331 | 331 |
| 12 | 322 | 266 | 273 | 343 | 313 | 313 |
| 13 | 395 | 317 | 339 | 414 | 385 | 389 |
| 14 | 378 | 256 | 274 | 402 | 326 | 339 |
| 15 | 309 | 235 | 244 | 333 | 313 | 310 |
| 16 | 275 | 248 | 243 | 314 | 260 | 263 |
| 17 | 292 | 239 | 242 | 296 | 288 | 280 |
| 18 | 327 | 317 | 297 | 364 | 309 | 316 |
| 19 | 263 | 231 | 226 | 314 | 275 | 275 |
| 20 | 259 | 239 | 230 | 278 | 247 | 240 |
| 21 | 395 | 303 | 330 | 402 | 338 | 352 |
| 22 | 327 | 254 | 264 | 314 | 322 | 310 |
| 23 | 378 | 317 | 330 | 353 | 317 | 319 |
| 24 | 378 | 249 | 268 | 370 | 307 | 316 |
| 25 | 332 | 293 | 293 | 377 | 348 | 348 |
| 26 | 313 | 251 | 257 | 358 | 267 | 283 |
| 27 | 395 | 303 | 330 | 464 | 363 | 397 |
| 28 | 378 | 280 | 305 | 402 | 334 | 348 |
| 29 | 354 | 249 | 265 | 338 | 331 | 324 |
| 30 | 322 | 293 | 287 | 393 | 334 | 345 |
| 31 | 287 | 253 | 251 | 348 | 338 | 331 |
| 32 | 346 | 303 | 305 | 393 | 242 | 310 |
| 33 | 317 | 270 | 274 | 384 | 326 | 335 |
| 34 | 292 | 286 | 271 | 377 | 311 | 321 |
| 35 | 339 | 293 | 297 | 384 | 343 | 348 |
| 36 | 271 | 249 | 243 | 358 | 249 | 270 |
| 37 | 354 | 286 | 301 | 393 | 331 | 342 |
| 38 | 317 | 270 | 274 | 370 | 328 | 333 |
| 39 | 317 | 303 | 287 | 348 | 317 | 318 |
| 40 | 346 | 317 | 310 | 358 | 324 | 325 |
| 41 | 313 | 273 | 274 | 358 | 292 | 303 |
| 42 | 309 | 303 | 282 | 364 | 324 | 327 |
| 43 | 309 | 256 | 259 | 348 | 319 | 319 |
| 44 | 313 | 280 | 278 | 338 | 284 | 291 |
| 45 | 378 | 303 | 323 | 444 | 290 | 315 |
| 46 | 309 | 265 | 267 | 348 | 284 | 294 |
| 47 | 313 | 276 | 276 | 348 | 331 | 327 |
| 48 | 395 | 317 | 339 | 488 | 348 | 389 |
| 49 | 263 | 265 | 252 | 305 | 251 | 255 |
| 50 | 327 | 266 | 274 | 343 | 301 | 305 |

| | | | | | | |
|----|-----|-----|-----|-----|-----|-----|
| 51 | 354 | 317 | 316 | 377 | 338 | 342 |
| 52 | 317 | 258 | 264 | 370 | 328 | 333 |
| 53 | 300 | 270 | 268 | 328 | 297 | 297 |
| 54 | 378 | 303 | 323 | 377 | 334 | 339 |
| 55 | 317 | 303 | 287 | 393 | 328 | 339 |
| 56 | 254 | 190 | 205 | 255 | 284 | 263 |
| 57 | 283 | 244 | 243 | 264 | 307 | 315 |
| 58 | 346 | 317 | 310 | 370 | 286 | 302 |
| 59 | 354 | 293 | 305 | 402 | 385 | 382 |
| 60 | 354 | 254 | 269 | 353 | 313 | 316 |
| 61 | 395 | 317 | 339 | 488 | 385 | 426 |
| 62 | 309 | 303 | 282 | 343 | 297 | 302 |
| 63 | 279 | 263 | 256 | 271 | 277 | 262 |
| 64 | 346 | 303 | 305 | 393 | 348 | 355 |

LIST OF REFERENCES

1. Lord, F. M., The Measurement of Growth, Educational and Psychological Measurement, Vol. 16:4, Winter 1956.
2. Lord, F. M., Large-Sample Covariance Analysis When the Control Variable is Fallible, American Statistical Association Journal, June 1960.
3. Lord, F. M., Further Problems in the Measurement of Growth, Educational and Psychological Measurement, Vol. 18:3, Autumn 1958.
4. McNemar, Q., On Growth Measurement, Educational and Psychological Measurement, Vol. 18:1, Spring 1958.
5. Cronbach, L. J. and Furby, L., How Should We Measure "Change" - or Should We? Psychological Bulletin, Vol. 74:1, July 1970.
6. Harris, C. W. (ed.), Problems in Measuring Change, The University of Wisconsin Press, Madison, Wisconsin, 1967.
7. California Achievement Tests, Examiners Manual Level 1A and 2A, CTB/McGraw-Hill, Monterey, California, 1970.
8. California Achievement Tests, Bulletin of Technical Data-Number 1, CTB/McGraw-Hill, Monterey, California, 1970.
9. Technical Report, California Achievement Test/1970 Edition, CTB/McGraw-Hill, Monterey, California, unpublished.
10. Gulliksen, H., Theory of Mental Tests, John Wiley and Sons, Inc., New York, 1950.
11. McNemar, Q., Psychological Statistics, 4th ed., John Wiley and Sons, Inc., New York, 1969.
12. Scott, W. A., and Wertheimer, M., Introduction to Psychological Research, 3rd ed., John Wiley and Sons, Inc., New York, 1966.
13. Guenther, W. C., Analysis of Variance, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1964.
14. Feldt, L. S., A Comparison of the Precision of Three Experimental Designs Employing a Concomitant Variable, Psychometrika, Vol. 23:4, December 1958.

15. Siegel, S., Non-Parametric Statistics, McGraw-Hill, New York, 1956.
16. McRae, D. J., MIKCA: A Fortran IV Iterative K-Mean Cluster Analysis Program, Behavioral Sciences, July 1971.
17. Gage, N. L., (ed.), Handbook of Research on Teaching, Rand McNally and Company, 1963.

INITIAL DISTRIBUTION LIST

| | No. Copies |
|---|------------|
| 1. Defense Documentation Center Cameron Station Alexandria, Virginia 22314 | 2 |
| 2. Library, Code 0212 Naval Postgraduate School Monterey, California 93940 | 2 |
| 3. Professor R. R. Read, Code 55 Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940 | 1 |
| 4. Chief of Naval Personnel Pers 11-b Department of the Navy Washington, D. C. 20370 | 1 |
| 5. Naval Postgraduate School Department of Operations Research and Administrative Sciences Monterey, California 93940 | 1 |
| 6. Mr. D. J. McRae CTB/McGraw-Hill Del Monte Research Park Monterey, California 93940 | 1 |
| 7. Dr. Henry T. Hutchins Assistant Superintendent Elementary Education Monterey Peninsula Unified School District P. O. Box 1031 Monterey, California 93940 | 2 |
| 8. CAPT T. T. Bean, USMC 24506 Chrisanta Mission Viejo, California 92675 | 2 |

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

| | | | |
|--|--|---|-----------------------|
| 1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940 | | 2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED | |
| | | 2b. GROUP | |
| 3. REPORT TITLE AN EXPERIMENT IN PARTITIONING SECOND-GRADE MATHEMATICS CLASSES INTO SMALLER SECTIONS OF COMPARABLE ABILITY | | | |
| 4. DESCRIPTIVE NOTES (Type of report and, inclusive dates) Master's Thesis; September 1971 | | | |
| 5. AUTHOR(S) (First name, middle initial, last name) Theodore T. Bean | | | |
| 6. REPORT DATE September 1971 | | 7a. TOTAL NO. OF PAGES 53 | 7b. NO. OF REFS 17 |
| 8a. CONTRACT OR GRANT NO. | | 9a. ORIGINATOR'S REPORT NUMBER(S) | |
| b. PROJECT NO. | | | |
| c. | | 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) | |
| d. | | | |
| 10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited. | | | |
| 11. SUPPLEMENTARY NOTES | | 12. SPONSORING MILITARY ACTIVITY Naval Postgraduate School Monterey, California 93940 | |
| 13. ABSTRACT <p>The mathematics program of an elementary school was changed so that the classes were partitioned into smaller sections of comparable ability. This change reduced the pupil-teacher ratio and the length of the instruction period. The proposal produced an unanswered question: Will there be an increased achievement gain during the school year which can be attributed to the program change? Two schools served as control groups for the experiment. The California Achievement Test, 1970 Edition, was used in a pretest-posttest design to measure achievement levels before and after the elapsed time of the experiment. The question was answered in the affirmative through the use of various statistical techniques: randomized matched subjects design, analysis of covariance and the non-parametric Mann-Whitney U test. The several techniques were applied since no single standard practice was available for this problem.</p> | | | |

| KEY WORDS | LINK A | | LINK B | | LINK C | |
|--------------------------------------|--------|----|--------|----|--------|----|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| Measuring and Evaluating Change | | | | | | |
| School Characteristics Questionnaire | | | | | | |
| Pretest-Posttest Design | | | | | | |
| California Achievement Test | | | | | | |
| Randomized Matched Subject Design | | | | | | |
| Analysis of Variance | | | | | | |
| Mann-Whitney U test | | | | | | |
| Cluster Analysis Computer Program | | | | | | |

1 JUN 76
5 SEP 78

23354
24807

Thesis

130552

B288

Bean

c.1

An experiment in partitioning second-grade mathematics classes into smaller sections of comparable ability.

1 JUN 76
5 SEP 78

23354
24807

Thesis

130552

B288

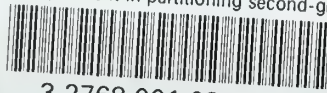
Bean

c.1

An experiment in partitioning second-grade mathematics classes into smaller sections of comparable ability.

thesB288

An experiment in partitioning second-gra



3 2768 001 03456 4

DUDLEY KNOX LIBRARY